Statistics: Lecture 1 - Probability Theory and Random Variables

Jan Bauer

jan.bauer@dhbw-mannheim.de

18.11.19

(Lukacs 1.2, 1.3, 2.1, 2.2, 3 & Soong 3.1)

Table of Contents

Table of Contents

Vocabs

- simple events/outcomes $\omega_1, \omega_2, \ldots$: Relevant outcomes for a random phenomena
- sample space Ω: Set of all simple outcomes
- (random) experiment: Procedure having a defined set of possible outcomes
- event: Subset of the sample space (all events: All subsets)
- An event is said to occur when one of the outcomes it contains occurs
- complementary event to the event E_i: Set of outcomes not contained in E_i (denoted by E_i^c)

Example: Coin Toss

- experiment: Tossing a coin
- simple outcomes: $\omega_1 = H$ (head) and $\omega_2 = T$ (tail)
- sample space:

$$\Omega = \{H, T\}$$

- events: $E_1 = H$ and $E_2 = T$

Example: Tossing a Coin twice

- experiment: Tossing a coin twice

- basic outcomes: $\omega_1 = HH$, $\omega_2 = HT$, $\omega_3 = TH$ and $\omega_4 = TT$,

- sample space:

$$\Omega = \{HH, HT, TH, TT\}$$

- (possible) events:

  $E_1 = \{HH, HT, TH\}$ ("at least one head")

  $E_2 = \{HH, TT\}$ ("both faces same")

  $E_1^c = \{TT\}$ ("no head") (complement of E_1)

Vocabs (classic algebra of sets)

- certain event: Ω
- impossible event: $\emptyset$
- union of A and B: Set of sample points that belong to at least A or B (denoted by $A \cup B$)
- intersection of A and B: Set of sample points that belong to both A and B (denoted by $A \cap B$)

Example (previous)

- $E_1 \cup E_2 = \Omega$
- $E_1 \cap E_2 = \{HH\}$

Probability

- Probability is a value assigned to a simple outcome.

  Consider $\Omega = \{\omega_1, \ldots, \omega_n\}$. To each ω_i, we assign its probability $\mathrm{P}(\omega_i) = p_i$, which satisfies

$$p_i \in [0,1] \text{ and } \sum_{i=1}^{n} p_i = 1 \ .$$

  For an event A, we assign its probability

$$\mathrm{P}(A) = \sum_{\omega_i \in A} p_i \ .$$

  $\rightsquigarrow$ the probability of an event A is the sum of the probabilities of the simple outcomes contained in this event.

Probability (Example)

Tossing a coin twice:

- $\mathrm{P}(\omega_1) = \mathrm{P}(HH) = p_1 = \frac{1}{4}$,
  $\mathrm{P}(\omega_2) = \mathrm{P}(HT) = p_2 = \frac{1}{4}$,
  $\mathrm{P}(\omega_3) = \mathrm{P}(TH) = p_3 = \frac{1}{4}$,
  $\mathrm{P}(\omega_4) = \mathrm{P}(TT) = p_4 = \frac{1}{4}$ with

$$\sum_{i=1}^{4} p_i = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1 \quad .$$

- $E_1 = \{HH, HT, TH\}$ ("at least one head")

$$\rightsquigarrow \mathrm{P}(E_1) = \mathrm{p}(HH) + \mathrm{p}(HT) + \mathrm{p}(TH) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

- $E_1^c = \{TT\}$ ("no head")

$$\rightsquigarrow \mathrm{P}(E_1^c) = 1 - P(E_1) = 1 - \frac{3}{4} = \frac{1}{4}$$

(probability of a complementary event: $\mathrm{P}(E) = 1 - \mathrm{P}(E^c)$)

DHBW Duale Hochschule Baden-Württemberg Mannheim

Kolmogorov Axioms

Let $\mathcal{E}$ be the event space containing all events. For the events $E_1, E_2, \dots$ in $\mathcal{E}$ and the sample space Ω it holds that

- non-negativity:

$$\mathrm{P}(E_i) \geq 0 \qquad \forall E_i \in \mathcal{E}$$

- unit measure:

$$\mathrm{P}(\Omega) = 1$$

- σ-additivity: If $E_1, E_2, \dots$ are disjoint ($E_i \cap E_j = \emptyset$, $j \neq i$), then

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(E_i)$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Consequences of Kolmogorov's Axioms

For the events $E_1, E_2, \ldots$ in the event space $\mathcal{E}$ and the sample space Ω it holds that

- $\mathrm{P}(\emptyset) = 0$
- if $E_i \cap E_j = \emptyset \;\Rightarrow\; \mathrm{P}(E_i \cap E_j) = 0$
- $\mathrm{P}(E_i) \in [0, 1] \qquad \forall E_i \in \mathcal{E}$
- $\mathrm{P}(E_i) = 1 - \mathrm{P}(E_i^c) \qquad \forall E_i \in \mathcal{E}$
- if $E_i \subseteq E_j \;\Rightarrow\; \mathrm{P}(E_i) \leq \mathrm{P}(E_j)$

The Addition Rule

- For any two events A and B, it holds that

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \quad .$$

- For the events $E_1, E_2, \ldots$, it holds that

$$\mathrm{P}\left(\bigcup_i E_i\right) \leq \sum_i \mathrm{P}(E_i) \quad .$$

(note the difference to σ-additivity)

The Addition Rule (Example)

Example: Roll a die once. What is the probability, that the number is even or that the number is smaller than four?

$\rightsquigarrow$ $A = \{2, 4, 6\}$ and $B = \{1, 2, 3\}$. Then

$$
\begin{aligned}
& \mathrm{P}(A \cup B) \\
= \; & \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \\
= \; & \mathrm{P}(\{2, 4, 6\}) + \mathrm{P}(\{1, 2, 3\}) - \mathrm{P}(\{2\}) \\
= \; & \frac{3}{6} + \frac{3}{6} - \frac{1}{6} \\
= \; & \frac{5}{6}
\end{aligned}
$$

Probability Assignment: Probability Trees

- Used to represent a series of events.
- Example: Tossing a coin twice. A is the first toss, B the second one.

Probability Assignment: Probability Tables (two Events)

Example: Tossing a coin twice. A is the first toss, B is the second toss.

Then the probability table providing the joint probabilities is given by

	$A = H$	$A = T$	$\mathrm{P}(B)$
$B = H$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$B = T$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$\mathrm{P}(A)$	$\frac{1}{2}$	$\frac{1}{2}$	1

- $\mathrm{P}(A = H \cap B = H) = \frac{1}{4}$
- $$\begin{aligned} & \mathrm{P}(\{A = H \cap B = H\} \cup \{A = T \cap B = T\}) \\ = \; & \mathrm{P}(A = H \cap B = H) + \mathrm{P}(A = T \cap B = T) \\ = \; & \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

  since $\{A = H \cap B = H\} \cap \{A = T \cap B = T\} = \emptyset$ are disjoint sets

DHBW Duale Hochschule Baden-Württemberg Mannheim

Example: Tossing a coin three times. A is the first toss, B is the second toss and C is the third toss, respectively. Then the probability table providing the joint probabilities is given by

$\boldsymbol{C = H}$	$A = H$	$A = T$	$\mathrm{P}(B \cap C = H)$
$B = H$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
$B = T$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
$\mathrm{P}(A \cap C = H)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

$\boldsymbol{C = T}$	$A = H$	$A = T$	$\mathrm{P}(B \cap C = T)$
$B = H$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
$B = T$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
$\mathrm{P}(A \cap C = T)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

Conditional Probability

- What is the probability of me going to a party tonight?

- What is the probability of me going to party tonight, when there is beer for free?

Conditional Probability

- Given two events A and B, the conditional probability of A given B is defined as

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \quad , \mathrm{P}(B) > 0 \quad .$$

- Equivalently, it holds that

$$\mathrm{P}(A \cap B) = \mathrm{P}(A|B) \cdot \mathrm{P}(B) \quad .$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Conditional Probability (Example)

Consider the event P with outcomes y and n, which corresponds to the outcome of going to the party (yay) or not (nay). Further, we have the event of beer for free, B, which has the outcomes y and n. We assume the following probability table:

	$P = y$	$P = n$	$\mathrm{P}(B)$
$B = y$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{5}{8}$
$B = n$	0	$\frac{3}{8}$	$\frac{3}{8}$
$\mathrm{P}(P)$	$\frac{1}{2}$	$\frac{1}{2}$	1

What is the probability of me going to party tonight, when there is beer for free?

$$\mathrm{P}(P = y | B = y) = \frac{\mathrm{P}(P = y \cap B = y)}{\mathrm{P}(B = y)} = \frac{1/2}{5/8} = \frac{4}{5} \,.$$

The Total Probability Rule

Let $E_1, E_2, \ldots$ be disjoint events with $\bigcup_i E_i = \Omega$ and suppose that $\mathrm{P}(E_i) > 0 \; \forall i$. Let A be an arbitrary event, then

$$\mathrm{P}(A) = \sum_i \mathrm{P}(E_i)\mathrm{P}(A|E_i) \; .$$

Bayes' Theorem

Let $E_1, E_2, \ldots$ be events and suppose that $\mathrm{P}(E_i) > 0 \; \forall i$. Let A be an arbitrary event satisfying $\mathrm{P}(A) > 0$, then

$$\mathrm{P}(E_i|A) = \frac{\mathrm{P}(E_i)\mathrm{P}(A|E_i)}{\mathrm{P}(A)} \overset{(TPR)}{=} \frac{\mathrm{P}(E_i)\mathrm{P}(A|E_i)}{\sum_i \mathrm{P}(E_i)\mathrm{P}(A|E_i)} \, ,$$

where the last conclusion follows if the Total Probability Rule assumptions

are satisfied.

Independence

- What is the probability of me going to a party tonight?
- What is the probability of me going to party tonight, when there is bad weather in England?

Independence

- Two events A and B are said to be independent, if

$$\mathrm{P}(A|B) = \mathrm{P}(A) \quad \Big(\text{or } \mathrm{P}(B|A) = \mathrm{P}(B)\Big) \quad .$$

- Equivalently, it holds that

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B) \quad .$$

- Where does this come from?

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A) \;\Leftrightarrow\; \mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B) \quad .$$

Independence of more than two events

- The events $E_1, E_2, \ldots$ are said to be independent, if

$$P\left(\bigcap_i E_i\right) = \prod_i P(E_i) \quad .$$

Check For Independence (Example)

Consider again the following probability table:

	$P = y$	$P = n$	$\mathrm{P}(B)$
$B = y$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{5}{8}$
$B = n$	0	$\frac{3}{8}$	$\frac{3}{8}$
$\mathrm{P}(P)$	$\frac{1}{2}$	$\frac{1}{2}$	1

Are P and B independent? Check $\mathrm{P}(P \cap B) = \mathrm{P}(P) \cdot \mathrm{P}(B)$ for all outcomes y and n:

$$\mathrm{P}(P = y \cap B = y) = \frac{1}{2} \neq \frac{5}{16} = \mathrm{P}(P = y) \cdot \mathrm{P}(B = y)$$

$$\rightsquigarrow \text{ not independent}$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Table of Contents

Table of Contents

Random Variables (Introduction)

- We want to consider random experiments associated with numerical outcomes

- Example: Rolling a die $\rightsquigarrow$ outcomes $1, 2, 3, 4, 5$ and 6

- Example: Tossing a coin $\rightsquigarrow$ translate outcomes "H" and "T" to 1 and 0

- Random variables are the base of probability theory

  We will only focus on the application, not the theory

- The definitions are not always mathematically complete; although sufficient for this course

Random Variable (definition)

- Our (sufficient) definition of a random variable (rv):

  A random variable X is a function

  $$X : \Omega \to \mathbb{R}, \;\; \omega \mapsto x = X(\omega)$$

  from a sample space, Ω, to the real numbers, $\mathbb{R}$.

- X is called a discrete rv if it is defined over a sample space having finite or a countably infinite number of sample points. In this case, X takes on discrete values.

- X is called a continuous rv if it is distributed over one or more continuous intervals in $\mathbb{R}$.

Random Variable (Example: Coin Toss)

- sample space:

$$\Omega = \{H, T\}$$

- we want to build a RV X that assigns the value 1 for "Head" and 0 for "Tail":

$$X(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{if } \omega = T \end{cases}$$

Random Variable (Example: Rolling a die)

- sample space:

$$\Omega = \{"1", "2", "3", "4", "5", "6"\}$$

- we want to build a RV X that assigns the value 1 when the die shows "1", the value 2 when the die shows "2",...

$$X(\omega) = \begin{cases} 1 & \text{if } \omega = "1" \\ 2 & \text{if } \omega = "2" \\ 3 & \text{if } \omega = "3" \\ \vdots & \vdots \end{cases}$$

Random Variable (Example: Rolling a die "reverse")

- sample space:

$$\Omega = \{"1", "2", "3", "4", "5", "6"\}$$

- we want to build a RV X that assigns the value 6 when the die shows "1", the value 5 when the die shows "2",...

$$X(\omega) = \begin{cases} 6 & \text{if } \omega = "1" \\ 5 & \text{if } \omega = "2" \\ 4 & \text{if } \omega = "3" \\ \vdots & \vdots \end{cases}$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Random Variable (Properties)

Let X and Y be two random variables. Then

- $X + Y$ is a RV
- $X - Y$ is a RV
- $c \cdot X$ is a RV (for $c \in \mathbb{R}$ a constant)
- X^2 is a RV
- $X \cdot Y$ is a RV
- $\frac{1}{X}$ is a RV (if $X \neq 0$)
- $\frac{X}{Y}$ is a RV (if $Y \neq 0$)